2.9 Systems of Linear Equations

This section is a review of solving systems of linear eq. We will see how the theory we have developed can be applied.

2.9.1 Definition

A system of m linear eq. in n variables $x_1, \dots, x_n$ over a field F is of the form

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m \qquad (*) \end{aligned}$$

for some $a_{ij}, b_k \in F$. A solution of $(*)$ is a tuple $(x_1, \dots, x_n) \in F^n$ which satisfies all m equations. Two systems over F of the same size are called equivalent if they have the same set of solutions. □

Questions: Given a system, does it have a solution? If yes, how many? Describe the set of solutions.

2.9.2 Definition An **elementary operation** on a system of linear equations over F is one of the following:

(a) swap two equations

(b) multiply an equation by a nonzero scalar

(c) add a nonzero multiple of one eq. to another. $\square$

2.9.3 Proposition Any two systems related by a sequence of elementary operations are equivalent. $\square$

2.9.4 Definition Associated to the system $(*)$ in 2.9.1 are the following matrices

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \qquad x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad b=\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \qquad (A|b)=\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

coefficient matrix **variable vector** **constant vector** **augmented matrix**

The **matrix form** of $(*)$ is the matrix equation $Ax=b$. The system $(*)$ is **homogeneous** if $b=0$. $\square$

2.9.5 Definition An **elementary row operation** on a matrix $A \in F^{m \times n}$ is one of the following

(a) swap two rows $R_i \longleftrightarrow R_j$, $i \neq j$

(b) multiply a row by a nonzero scalar $R_i \leftarrow \alpha R_j, \alpha \neq 0$

(c) add a nonzero mult of one row to another $R_i \leftarrow R_i + \alpha R_j, \alpha \neq 0, i \neq j$

Two matrices are row equivalent if they are related by a sequence of elementary row operations.

$\square$

2.9.6 Proposition Row equivalence is an equivalence relation on $F^{m \times n}$.

$\square$

2.9.7 Corollary Row equivalent augmented matrices represent equivalent systems of equations.

Proof. Translate Prop 2.9.3 into matrix form

$\square$

The point is that we can translate everything into matrices. Then one solves the system by row reducing the augmented matrix until it is in a simpler form.

2.9.8 Definition A matrix $A \in F^{m \times n}$ is in reduced row echelon form if

① All zero rows (if any) are at the bottom.

② The leading coeff of any nonzero row is strictly to the right of the leading coefficient of the row above it.

③ the leading coefficient of every nonzero row is 1

④ the leading coefficient of every nonzero row is the only nonzero entry in its column.

$\square$

2.9.9 Example

Reduced row echelon form: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$\begin{pmatrix} 0 & 1 & 0 & -1 \end{pmatrix}$

Not reduced row echelon form: $\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$\square$

Recall: There exists an algorithm called Gauss-Jordan reduction, that takes in a matrix and produces a matrix in reduced row echelon form using only elementary row operations.

2.9.10 Proposition

Every matrix $A \in F^{m\times n}$ is row equivalent to a unique matrix in reduced row echelon form. The unique such matrix is denoted by $\mathrm{Rref}(A)$.

Proof. Omitted. $\square$

Now, lets focus on homogeneous systems $Ax=0$. The solution space is the kernel of the linear map

$$F^n \xrightarrow{A\cdot} F^m \qquad A \in F^{m\times n}$$

So, the system always has the trivial solution $x=0$, and has a nontrivial solution iff $\dim(\ker(A)) = \text{nullity}(A) > 0$.

By Rank/Nullity

$$\dim(\text{sol'n space}) = \text{nullity}(A)$$
$$= n - \text{rank}(A).$$

So, the system has a nontrivial solution if $n > \text{rank}(A)$. For this reason, one is interested in the following spaces:

(row space) $\text{Row}(A) = \text{span}\{\text{rows of } A\} \subseteq F^n$

$$= \text{im}\left(F^m \xrightarrow{A^T\cdot} F^n\right)$$

(column space) $\text{Col}(A) = \text{span}\{\text{columns of } A\} \subseteq F^n$

$$= \text{im}\left(F^n \xrightarrow{A\cdot} F^m\right)$$

Notice: $\text{rank}(A) = \dim(\text{im}(A\cdot))$

$$= \dim(\text{col}(A)).$$

$=$ maximum # of independent columns of A.

Performing row operations on A amounts to post-composing $F^n \xrightarrow{A\cdot} F^m$ w/ certain isomorphisms $F^m \xrightarrow{\sim} F^m$.

The isomorphisms are given by multiplication by "elementary matrices":

2.9.11 Definition The **elementary matrices** are the square matrices of the following form.

(a) $I_m + \lambda E_{ij}$, $i \neq j$, $\lambda = 0$

$$\begin{array}{cc} & \begin{array}{ccc} & & j \end{array} \\ i & \begin{pmatrix} 1 & & \lambda \\ & \ddots & \\ & & 1 \end{pmatrix} \end{array}$$

(b) $I_m - E_{ii} - E_{jj} + E_{ij} + E_{ji}$, $i \neq j$

$$\begin{array}{cc} & \begin{array}{cccccc} & & i & & j & \end{array} \\ \begin{array}{c} \\ i \\ j \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & & 1 & \\ & & & \ddots & & \\ & & 1 & & 0 & \\ & & & & & \ddots \; 1 \end{pmatrix} \end{array}$$

(c) $I_m + (\lambda - 1) E_{ii}$

$$\begin{array}{cc} & \begin{array}{ccc} & i & \end{array} \\ i & \begin{pmatrix} 1 & & \\ & \ddots \; \lambda & \\ & & \ddots \; 1 \end{pmatrix} \end{array}$$

☐

Observe:

(a) $(I_n + \lambda E_{ij})A = A + \lambda \boxed{E_{ij}A}$

j^{th} row of A placed in row i

⇝ Adds a λ multiple of jth row to i^{th} row

$(R_i \longleftarrow R_i + \lambda R_j)$

(b) $(I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji})A$

$$= A - \underbrace{E_{ii}A}_{\text{removes } i\text{th row}} - \underbrace{E_{jj}A}_{\text{removes } j\text{th row}} + \underbrace{E_{ij}A}_{\text{adds } j\text{th row to } i\text{th row}} + \underbrace{E_{ji}}_{\text{adds } i\text{th row to } j\text{th row}}$$

⇝ swaps rows i and j $(R_i \longleftrightarrow R_j)$

(c) Similarly, multiplies i^{th} row by λ $(R_i \longleftarrow \lambda R_i)$

Elementary matrices are invertible since the corresponding row operations are. Thus, A, B are row equivalent iff there is an invertible matrix E such that $B = EA$.

2.9.12 Proposition If R is in reduced row echelon form, then the nonzero rows form a basis for $\text{Row}(R)$ and the leading columns form a basis for $\text{Col}(R)$. □

2.9.13 Proposition Let $A = (A_1 \cdots A_n)$ and $B = (B_1 \cdots B_n)$ be row equivalent matrices. The columns $A_{j_1}, \dots, A_{j_r}$ are independent in F^m if and only if the columns $B_{j_1}, \dots, B_{j_r}$ are independent in F^m.

Proof. $B = EA$ for invertible matrix E. So $B_j = EA_j$ for all j. Then

$$\sum_{i=1}^{n} \alpha_i B_{j_i} = \sum_{i=1}^{n} \alpha_i E A_{j_i} = E \sum_{i=1}^{n} \alpha_i A_{j_i}.$$

Thus, $\sum_{i=1}^{n} \alpha_i B_{j_i} = 0$ if and only if $\sum_{i=1}^{n} \alpha_i A_{j_i}$ since E is invertible.

□

2.9.14 Proposition The columns of A corresponding to the leading columns of $\text{Rref}(A)$ form a basis for $\text{Col}(A)$.

Proof. Prop 2.9.12 + Prop 2.9.13.

□

2.9.15 Proposition If A and B are row equivalent, then $\mathrm{Row}(A) = \mathrm{Row}(B)$.

Proof. $B = EA$ for some invertible matrix. Then

$B^T = A^T E^T$.

$$\begin{tikzcd} F^m \arrow[rr, "B^T", "\circlearrowleft"'] \arrow[dr, "E^T \wr"'] & & F^n \\ & F^m \arrow[ur, "A^T"'] & \end{tikzcd}$$

Precomposition by an isomorphism doesn't change the image.

Thus, $\mathrm{Row}(B) = \mathrm{im}(B^T) = \mathrm{im}(A^T E^T) = \mathrm{im}(A^T) = \mathrm{Row}(A)$. □

2.9.16 Proposition The nonzero rows of $\mathrm{Rref}(A)$ form a basis for $\mathrm{Row}(A)$.

Proof. A and $\mathrm{Rref}(A)$ are row equivalent, so they have the same row space by 2.9.15. By 2.9.12, the nonzero rows of $\mathrm{Rref}(A)$ are a basis.

□

2.9.17 Corollary $\mathrm{rank}(A) = \dim(\mathrm{col}(A)) = \dim(\mathrm{row}(A))$.

Proof.

$$\begin{aligned} \dim(\mathrm{row}(A)) &= \dim(\mathrm{row}(\mathrm{Rref}(A))) \\ &= \text{number of nonzero rows of } \mathrm{Rref}(A) \\ &= \text{number of leading entries of a row} \\ &= \text{number of leading columns} \\ &= \dim(\mathrm{col}(A)). \end{aligned}$$

□

Thus, $Ax = 0$ has a nontrivial solution if and only if $\dim(\mathrm{row}(A)) < n$. The point is that this can be computed using row operations. So for instance, if $A \in F^{m \times n}$ w/ $m<n$, then $Ax=0$ has a nontriv solution. If $m=n$, then $Ax=0$ has a nontrivial solution if and only if $\mathrm{Rref}(A)$ has a zero row. The methods can be extended to nonhomogeneous systems $Ax=b$. Row reduce $(A|b)$ to find a single sol'n x_0, $Ax_0=b$. Then if $x \in \ker(A)$, $x+x_0$ is also a solution: $A(x+x_0) = Ax + Ax_0 = 0 + b = b$. So the set of solutions to $Ax=b$ is $\ker(A)+x_0$.

Chapter III: Duality and Multilinear Algebra

3.1 Linear Functionals / Dual Space

3.1.1 Definition

A linear functional on a vector space V over F is just a linear map $f: V \to F$. ☐

3.1.2 Example

(a) The trace of a matrix defines a linear functional

$$\begin{aligned} \mathrm{tr}: F^{n\times n} &\longrightarrow F \\ A=(a_{ij}) &\longmapsto \sum_{i=1}^{n} a_{ii} \end{aligned}$$

(b) Evaluation of a polynomial at any $\alpha \in F$ define a linear function

$$ev_\alpha : F[x] \longrightarrow F$$
$$p(x) = \sum_{i=0}^{n} a_i x^i \longmapsto \sum_{i=0}^{n} a_i \alpha^i =: p(\alpha)$$
$$\alpha^0 := 1$$

(c) For any compact interval $[a,b] \subseteq \mathbb{R}$, integration defines a linear functional

continuous functions $f : [a,b] \to \mathbb{R}$ $=$
$$C([a,b]) \longrightarrow \mathbb{R}$$
$$f \longmapsto \int_a^b f\,dx.$$

$\square$

3.1.3 Definition The **dual space V^*** of a vector space V is the set of linear functionals on V:

$$V^* := \mathrm{Hom}_F(V,F). \quad \left[\begin{array}{l}(f+g)(v) = f(v)+g(v)\\ (\alpha f)(v) = \alpha f(v)\end{array}\right]$$

By Theorem 2.5.1, it is a vector space over F.

$\square$

Remark: Using Cor 2.6.3, we already know that if V is finite-dimensional, then

$$\begin{aligned}\dim_F V^* &= \dim_F \mathrm{Hom}(V,F)\\ &= \dim_F V \cdot \dim_F F = \dim_F V.\end{aligned}$$

So if V is finite-dimensional, then $V \cong V^*$. So the first step to understanding V^* is to find an isomorphism.

3.1.4 Theorem **(Dual Basis)** Let V be a finite-dimensional vector space over F w/ basis $B=(b_1,\dots,b_n)$. There exists a unique basis $B^*=(\varphi_1,\dots,\varphi_n)$ for V^* such that

$$\varphi_i(b_j)=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases} =: \delta_{ij} \qquad \text{"Kronecker delta"}$$

for all $i,j\in\{1,\dots,n\}$. The basis B^* is called the **dual basis** to B. Moreover, if $f\in V^*$, then

$$f=\sum_{i=1}^{n} f(b_i)\varphi_i \qquad \left(\Longleftrightarrow [f]_{B^*}=\left(f(b_1),\dots,f(b_n)\right)\right)$$

and if $v\in V$, then

$$v=\sum_{i=1}^{n}\varphi_i(v)\,b_i. \qquad \left(\Longleftrightarrow [v]_B=\left(\varphi_1(v),\dots,\varphi_n(v)\right)\right)$$

Proof. For each $i=1,\dots,n$, define $\varphi_i: V\longrightarrow F$ to be the unique linear map satisfying

$$\varphi_i(b_j)=\delta_{ij}.$$

$$\begin{tikzcd} b_j \arrow[r, mapsto] & \delta_{ij} \\ B \arrow[r] \arrow[d] & F \\ V \arrow[ur, dashed, "G", "\exists!"'] & \end{tikzcd}$$

Thm 2.3.1

By Theorem 2.3.1, the φ_i are linear maps uniquely determined by their values on the basis B. So it suffices to show that $B^*=(\varphi_1,\dots,\varphi_n)$ is a basis for V^*. Then the uniqueness of the basis is immediate.

(B^* is independent) Suppose $\sum_{i=1}^{n}\alpha_i\varphi_i=0$ (in V^*)

Then for each $j\in\{1,\dots,n\}$,

$$\begin{aligned} 0 = 0(b_j) = \left(\sum_i \alpha_i \varphi_i\right)(b_j) &= \sum_{i=1}^{n} \alpha_i \varphi_i(b_j) \\ &= \sum_{i=1}^{n} \alpha_i \delta_{ij} \\ &= \alpha_j . \end{aligned}$$

Thus, $\alpha_1 = \cdots = \alpha_n = 0$.

(B^* spans V^*) By dimension argument, B^* spans. But we have to prove the statements about coordinate vectors, which will follow from this part. Let $f \in V^*$. The claim is that

$$f = \sum_{i=1}^{n} f(b_i)\varphi_i .$$

Both sides are linear maps, so it suffices to show that both maps have the same values on B. Let $j \in \{1, \dots, n\}$. Then

$$\begin{aligned} \sum_{i=1}^{n} f(b_i)\varphi_i(b_j) &= \sum_{i=1}^{n} f(b_i)\delta_{ij} \\ &= f(b_j) . \end{aligned}$$

Lastly, let $v \in V$. The claim is that $v = \sum_{i=1}^{n} \varphi_i(v) b_i$.

Suppose $[v]_B = (\alpha_1, \dots, \alpha_n)$. We claim that $\alpha_i = \varphi_i(v)$. Let $i \in \{1, \dots, n\}$. Then

$$\begin{aligned} \varphi_i(v) = \varphi_i\left(\sum_{j=1}^{n} \alpha_j b_j\right) = \sum_{j=1}^{n} \alpha_j \varphi_i(b_j) &= \sum_{j=1}^{n} \alpha_j \delta_{ij} \\ &= \alpha_i . \end{aligned}$$

This proves the claim. $\square$

Remark: By the theorem, each basis of V gives rise to a basis of V^*. The linear functional f_i in the dual basis just gives the ith coordinate of $v \in V$ in the original basis for V.

Question: Does every basis for V^* arise as the dual basis of some basis of V? We can answer this question by considering the dual space $(V^*)^*$ of V^*. See Corollary 3.1.8.

3.1.5 Definition Let V be a vector space over F. For each $v \in V$, define $ev_v \in V^{**}$ as follows:

$$\begin{aligned} ev_v : V^* &\longrightarrow F \\ f &\longmapsto f(v) \end{aligned}$$

That is, ev_v is the function which evaluates a linear functional $f : V \to F$ at $v \in V$. $\square$

3.1.6 Theorem If V is a finite-dimensional vector space, then there is an isomorphism

$$\begin{aligned} V &\longrightarrow V^{**} \\ v &\longmapsto ev_v \end{aligned}$$

of vector spaces.

Proof. First, one checks that $ev_v \in V^{**}$. Exercise. Then one checks that $\begin{aligned} V &\longrightarrow V^{**} \\ v &\longmapsto ev_v \end{aligned}$ is linear.

Exercise. Then one checks that the map is an isomorphism. We already know $\dim V = \dim V^* = \dim V^{**}$, so it suffices to show that $V \longrightarrow V^{**}$ is injective, i.e. $\ker(v \longmapsto ev_v) = \{0\}$. Suppose $0 \neq v \in V$. We claim that $ev_v \neq 0$. Since $v \neq 0$, $\{v\}$ is independent. Extend $\{v\}$ to a basis for V. Define $f_v : V \longrightarrow F$ by specifying its values on the basis:

$$ev_v : V^* \longrightarrow F$$

$$f_v(v) = 1$$

$$f_v(w) = 0 \quad \text{for any basis vector } w \neq v.$$

Then $ev_v(f_v) = f_v(v) = 1$. Thus, $ev_v \neq 0$. By contrapositive, $ev_v = 0$ implies $v = 0$. That is $\ker(v \longmapsto ev_v) = \{0\}$. □

Remark: Under this identification $v \longmapsto ev_v$, one can think of V as being the dual space of V^*.

3.1.7 Corollary

Let V be a finite-dimensional vector space. If $L \in V^{**}$, then there is a unique $v \in V$ such that

$$L(f) = f(v). \quad \forall f \in V^*$$

In other words, linear functionals $L : V^* \longrightarrow F$ are just "evaluation" at different vectors in V.

Proof. Let $L \in V^{**}$. By Thm 3.1.6., there is a unique $v \in U$ such that $L = ev_v$. Then

$$L(f) = ev_v(f) = f(v).$$

□

3.1.8 Corollary Let V be a finite-dimensional vector space. Then every basis for V^* is the dual basis of a basis for V.

Proof. Let $C = (f_1, \dots, f_n)$ be a basis for V^*. Then $C^* = (g_1, \dots, g_n)$ is a basis for V^{**} satisfying

$$g_i(f_j) = \delta_{ij} \qquad (\text{Thm } 3.1.4)$$

Then using 3.1.7, there are unique vectors $b_1, \dots, b_n \in V$ such that $g_j(f) = f(b_j)$ for all $f \in V^*$. In particular,

$$f_i(b_j) = g_j(f_i) = \delta_{ji}.$$

So if $B = (b_1, \dots, b_n)$ is a basis for V, then $C = B^*$ by uniqueness in Thm 3.1.4. But B is a basis because $b_1, \dots, b_n$ are the images of the elements $g_1, \dots, g_n$ of the basis C^* for V^{**} under the isomorphism $V^{**} \longrightarrow V$. So B is a basis. $\square$